



id: SI_005

model: openai/gpt-5.4-mini

seeds:

04_risiko: usikkerhet

03_relasjon: makt

05_handling: vent

01_perspektiv: strategisk

02_tidhorisont: kortsiktig

Instans SI_005

Seeds: strategisk | kortsiktig | makt | usikkerhet | vent

Å samle **komplett prisdata på alle varer i verden** er i praksis ikke et dataprojekt først, men et **makt- og kontrollproblem**. Den strategiske feilen er å tenke “full dekning” som et statisk mål. Pris er ikke en egenskap ved varen alene; den oppstår i et bestemt marked, tidspunkt, pakningsstørrelse, valuta, skatt, kampanje og kjøpskanal.

Min lesning er kortsiktig og brutal: **vent på perfektjon, og du får ingenting**. Start med det som faktisk kan eies.

Hva som må bygges

1. Global vare-ontologi

- Bruk eksisterende standarder: **GTIN/EAN/UPC, GS1, UNSPSC**, eventuelt **HS-koder** for handel.
- En vare må normaliseres på nivåene: produktvariant, pakning, volum/vekt, region, kanal.

2. Prisinnsamlingsnettverk

- Egen innhenting fra nettbutikker, markedsplasser og offentlige priskilder.
- API-er der de finnes; ellers scraping der det er lovlig.
- Fysisk prisinnsamling for dagligvarer og apotek er nødvendig i mange markeder.

3. Maktkart

- Store aktører kontrollerer mye av tilgjengeligheten: **Amazon, Walmart, Alibaba, JD.com, Carrefour, Tesco**, samt lokale kjeder.



- De kan blokkere, rate-limite, endre struktur og skjule priser bak innlogging eller geografi. Du må anta motstand.

Strategi

- Ikke mål “alle varer”; mål **90/10 først**: de 10% av varetypene som står for 90% av prisvolumet i hver region.
- Bygg et system som tåler **usikkerhet**: manglende priser, midlertidige tilbud, geografiske forskjeller og substitutter.
- Bruk probabilistisk matching, ikke streng matching. “Komplett” blir en **konfidens-score**, ikke en binær tilstand.

Kortsiktig plan

- Velg 3 markeder: USA, EU, India.
- Velg 5 kategorier: dagligvarer, elektronikk, klær, legemidler, husholdning.
- Koble 100–200 datakilder først.

Neste steg

1. Definer en global produktmodell med GTIN + variant + pakningsenhet.
2. Bygg en pris-pipeline for 100 prioriterte kilder i tre land.
3. Innfør konfidensscore per pris, slik at “komplett” måles som dekning per kategori og marked.